

Periodically driven stochastic un- and refolding transitions of biopolymers

O. BRAUN¹ and U. SEIFERT¹

¹ *II. Institut für Theoretische Physik, Universität Stuttgart, 70550 Stuttgart, Germany*

PACS. 87.15.-v – Biomolecules: structure and physical properties.

PACS. 87.64.Dz – Scanning tunneling and atomic force microscopy.

Abstract. – Mechanical single molecule experiments probe the energy profile of biomolecules. We show that in the case of a profile with two minima (like folded/unfolded) periodic driving leads to a stochastic resonance-like phenomenon. We demonstrate that the analysis of such data can be used to extract four basic parameters of such a transition and discuss the statistical requirements of the data acquisition. As advantages of the proposed scheme, a polymeric linker is explicitly included and thermal fluctuations within each well need not to be resolved.

Force spectroscopy has been used extensively to probe mechanically the interactions within a single biomolecule, like RNA and DNA, peptides or proteins by applying an externally controlled load on such a structure; for reviews see [1–7]. Since in these experiments typically the extension along one coordinate r is controlled, the data and results are rationalized in terms of the notion of a (free) energy profile $G(r)$ along this coordinate. A major challenge for future improvements is to design these experiment in such a way that as much information as possible on such a profile can be extracted reliably from the data.

In the paradigmatic situation of an experiment using an AFM cantilever, the biomolecule is attached via a polymeric linker to the tip of the AFM whose base is driven according to a certain protocol, see fig. 1. Up to now, in most applications in force spectroscopy, the force is applied according to a linear ramp protocol $x(t) = x_0 + vt$. Here, x_0 is the cantilever position at time $t = 0$ and v is the constant ramp velocity. From the peak of the distribution of unfolding forces, one can extract two parameters of the underlying free energy landscape: the distance x_u from the first minimum to the maximum and the spontaneous off-rate k_{12}^0 [8], see fig. 2. These quantities characterize mainly the unfolding part of the profile. If one wants more information about the potential $G(r)$, in particular on quantities determining the refolding events, one has to resort to a more sophisticated analysis.

In the latest development, analyzing the distribution of the applied work spent in the unfolding process yields the full underlying profile using Jarzynski's relation [9–11]. In this context, we have shown that a periodic rather than a linear ramp improves the reconstruction of the profile significantly [12]. Such an analysis based on Jarzynski's relation, however, requires first that the position of the tip $z(t)$ is measured very accurately. Second, the reconstructed energy profile refers to the coordinate z which includes the linker molecule. It is nontrivial to extract from such a profile $G(z)$ the quantity $G(r)$ which is of main interest.

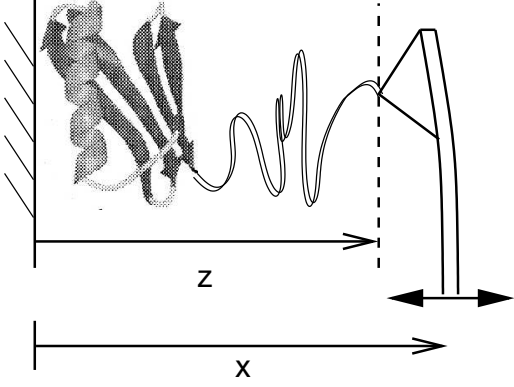


Fig. 1

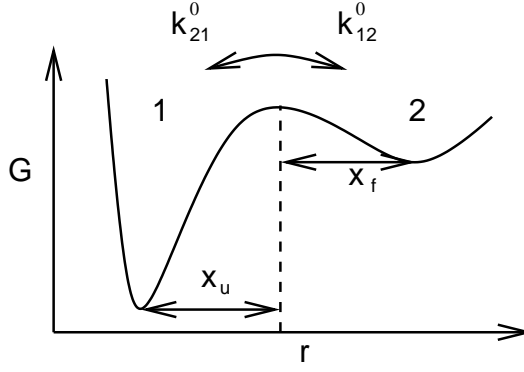


Fig. 2

Fig. 1 – Sketch of an atomic force spectroscopy experiment. The protein (thick lines) is connected via a polymeric linker (thin line) to the tip of an AFM cantilever at $z(t)$ whose base at $x(t)$ is externally controlled.

Fig. 2 – Schematic view of the underlying free energy potential $G(r)$ as a function of the reaction coordinate r . The first minimum represents the folded state, whereas the second shallow minimum represents the unfolded state of the biopolymer. x_u and x_f denote the respective distances from the minima to the barrier and k_{ij}^0 the spontaneous transition rates from state i to j .

In this letter, we propose a method on an intermediate level of sophistication, which yields, apart from x_u and k_{12}^0 , the corresponding two quantities x_f and k_{21}^0 of the refolding transition. We employ a periodic ramp which leads to a stochastic resonance-like phenomenon [13]. As advantages of our new scheme, we include the linker and do not require to analyze the full stochastic trajectory $z(t)$ as it is necessary in approaches based on Jarzynski's relation. It is sufficient to observe the sequence of folding and unfolding events reflected in this coordinate as explained below. Our proposal extends a previous approach where spontaneous transitions under constant force are used to extract these parameters [14]. The main advantage of our dynamical approach is that stochastic resonance leads to an enhancement of the transition. Moreover more control parameters such as amplitude and frequency of the driving are available to optimize the data acquisition. Finally no feed-back loop in order to impose a constant force is necessary.

We first recall that the spontaneous transition rates k_{12}^0 and k_{21}^0 for unfolding and refolding are modified by a force F according to [8, 15, 16]

$$k_{12} = k_{12}^0 \exp[\beta x_u F(z, L)] \quad (1)$$

$$k_{21} = k_{21}^0 \exp[-\beta x_f F(z, L)] \quad (2)$$

where β is the inverse temperature. The force $F(z, L)$ is transmitted from the cantilever to the biomolecule by a polymeric linker which we model by a worm-like chain as [17]

$$F(z, L) = \frac{1}{\beta L_p} \left(\frac{1}{4(1 - z/L)^2} - \frac{1}{4} + \frac{z}{L} \right), \quad (3)$$

where L_p is the persistence length. The contour length L depends on the configuration. In the folded state, it is $L = L_0$, in the unfolded state $L = L_0 + \Delta L$, where ΔL is the length gain of the contour length upon unfolding.

In each state, the force F_c measured by the cantilever with harmonic spring constant k_c has to balance the polymeric force

$$F_c \equiv k_c(x - z) = F(z, L) . \quad (4)$$

Thus, the z -coordinate can be calculated for given externally controlled cantilever base position x . In general eq. (4) can not be solved analytically for $z(x)$. In force spectroscopy on biopolymers, however, the linker is almost stretched out for a typical unfolding force of about 50-150 pN. Eqs. (3) and (4) then reduce to the cubic equation

$$k_c(x - L + \delta) = \frac{L^2}{4\beta L_p \delta^2} , \quad (5)$$

for $\delta \equiv L - z$, which denotes the length stored in the fluctuations. Hence the functions $z = z(x, L)$ or $\delta = \delta(x, L)$ also depend on the current length of the polymer.

We now apply a periodic protocol of the cantilever position x according to

$$x(t) = x_0 + x_a \cos(\omega t) , \quad (6)$$

with the offset extension x_0 , the amplitude x_a and frequency ω [12]. In general, the periodicity of $x(t)$ leads to an anharmonic periodicity of $\delta(t)$ via eq. (5), but the leading term for $x_a \ll x_0$ is given by the first harmonics as

$$\delta(t) \approx \delta_0 + \delta_a \cos(\omega t) , \quad (7)$$

where δ_0 is the solution of eq. (5) with $x = x_0$. The amplitude δ_a is given by

$$\delta_a = -\frac{x_a}{3 + 2(x_0 - L)/\delta_0} . \quad (8)$$

To first order in δ_a/δ_0 , the force acting on the biomolecule exhibits the same periodicity

$$F(\delta, L) \approx \frac{L^2}{4\beta L_p \delta_0^2} \left(1 - \frac{2\delta_a}{\delta_0} \cos(\omega t) \right) \equiv F_0 + F_a \cos(\omega t). \quad (9)$$

For the functions F_0 and F_a we will have to distinguish between the folded the unfolded configuration.

The probability $\Phi(t)$ that the molecule is in the folded state obeys the master equation

$$\frac{d}{dt}\Phi = -(k_{12} + k_{21})\Phi + k_{21}, \quad (10)$$

where the rates depend on the force (9) through (1) and (2). For small loading, an analytical solution can be calculated in linear order in x_a/x_0 by splitting the solution of the master equation (10) into a time-independent part Φ_0 and a time-dependent part $\Pi(t)$ as

$$\Phi(t) = \Phi_0 + \Pi(t) . \quad (11)$$

The stationary part is given by

$$\Phi_0 = \frac{\kappa_{21}}{\kappa_{12} + \kappa_{21}} , \quad (12)$$

with

$$\kappa_{12} \equiv k_{12}^0 \exp[\beta x_u F_0(\delta_0, L_0)] \quad \text{and} \quad \kappa_{21} \equiv k_{21}^0 \exp[-\beta x_f F_0(\delta_0, L_0 + \Delta L)] , \quad (13)$$

which are the time-independent contributions to the transition rates due to the preloading. On the intrinsic time scale $1/\Lambda \equiv (\kappa_{12} + \kappa_{21})^{-1}$, the time dependent part of $\Phi(t)$ approaches a limit cycle

$$\Pi(t) \equiv \Pi_0 \cos(\omega t - \lambda) = \frac{\alpha}{\sqrt{\omega^2 + \Lambda^2}} \cos(\omega t - \lambda) , \quad (14)$$

with

$$\alpha \equiv -(\beta x_u F_a(L_0) + \beta x_f F_a(L_0 + \Delta L)) \kappa_{12} \Phi_0 \quad (15)$$

and the phase shift

$$\lambda \equiv \arctan(\omega/\Lambda) . \quad (16)$$

The solution of eq. (10) is thus a periodic response of the probability of occupation Φ of the folded state.

For fixed x_0 and fixed ω the solution (11) yields three parameters α, Λ and Φ_0 for the four unknowns $x_u, x_f, k_{12}^0, k_{21}^0$. For a determination of these crucial parameters, more data are needed, which can easily be obtained by varying the offset extension x_0 and the frequency ω . Such a variation leads to an interesting resonance phenomenon.

It is one of the well known features of periodically driven stochastic transitions, that in contrast to a deterministic mechanical system, the response amplitude depends on the frequency in a non-resonant manner [13]. For low frequencies, where the stochastic transition can follow the externally driven loading, the amplitude is larger than for large frequencies, see eq. (14). Both, the response amplitude Π_0 of the periodic part of the occupation probability and the phase shift λ show a clear maximum as a function of the offset x_0 , see figs. 3 and 4. The maximal phase shift λ^* follows from the condition $\kappa'_{12} + \kappa'_{12} = 0$, where the prime denotes the derivative with respect to x_0 . This yields the ratio of the geometric parameters x_f and x_u as

$$\frac{x_f}{x_u} = \frac{\kappa_{12} F'_0(L_0)}{\kappa_{21} F'_0(L_0 + \Delta L)} , \quad (17)$$

where $F'_0(L_0)$ and $F'_0(L_0 + \Delta L)$ are the derivatives of the force with respect to x_0 with folded contour length and unfolded contour length respectively. With eq. (5), we get explicitly

$$F'_0(L) = \frac{L^2}{2\beta L_P \delta_0^{*2} (\delta_0^* + 2(\delta_0^* + x_0^* - L))} , \quad (18)$$

where x_0^* is the position of the maximum of the phase shift and $\delta_0^* = \delta_0(x_0^*, L)$. Eq. (17) then yields the ratio x_u/x_f of the underlying free energy profile as a fourth relation which can be used to determine $x_u, x_f, k_{12}^0, k_{21}^0$. Of course, many other fitting schemes can be envisaged to extract these crucial parameters.

Up to now, we have not yet addressed explicitly the question of how to obtain the probability $\Phi(t)$ from real measurements and how these experiments have to be designed to deliver the necessary data most efficiently. In principle, the occupation probability $\Phi(t)$ has to be measured via an ensemble average of trajectories. The ensemble average, however, can be

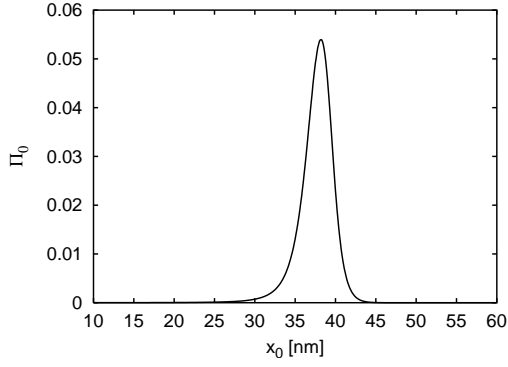


Fig. 3

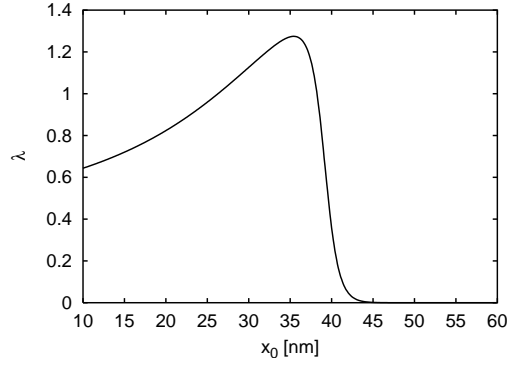


Fig. 4

Fig. 3 – The analytical solution of the response amplitude of the occupation probability of the folded state Π_0 as a function the offset extension x_0 . The parameters are $k_{21}^0 = 3 \text{ s}^{-1}$, $k_{12}^0 = 4.6 \cdot 10^{-4} \text{ s}^{-1}$, $x_u = 0.3 \text{ nm}$, $x_f = 1 \text{ nm}$, $L_p = 0.42 \text{ nm}$, $L_0 = 40 \text{ nm}$, $\Delta L = 28.5 \text{ nm}$, $\omega = 1 \text{ s}^{-1}$, $x_a = 0.5 \text{ nm}$, $k_c = 5 \text{ pN/nm}$, taken from [18] for a tenascin FN-III segment.

Fig. 4 – The phase shift λ of the occupation probability over the offset extension x_0 for the same parameters as in fig. 4.

replaced by a phase sensitive time average of a single trajectory, because $\Phi(t)$ is periodic with period time $T = 2\pi/\omega$ after a relaxation time for the offset expansion of the order of $t \sim 1/\Lambda$. Comparing these two averaging schemes, the advantage of the phase sensitive time average is obviously that the relaxation time has to be accounted for only once.

According to our assumptions, the trajectories $z(t)$ are bistable, so the conformation (folded or unfolded) of the biopolymer can be determined unambiguously from $z(t)$ since each transition leads to a jump in $z(t)$ which is larger than the thermal fluctuations within each minimum. We assign a weight $p(t) = 1$ if the molecule at time t is in the folded state and $p(t) = 0$ if it is in the unfolded state. The probability of occupation $\Phi(t)$ of the folded state can then be estimated by

$$\Phi_N(t) = \frac{\sum_{n=0}^{N-1} p_n(t)}{N}, \quad (19)$$

where $N = T_{\text{tot}}/T$ is the number of periods with the total measuring time T_{tot} and $p_n(t) \equiv p(t + nT)$ and $0 < t < T$. We now optimize ω and T_{tot} for the least statistical error.

The autocorrelation of $p(t)$ up to first order in x_a/x_0 is given by [13]

$$\text{cov}(p(t), p(t + \tau)) \equiv (\langle p(t) \rangle - p(t))(\langle p(t + \tau) \rangle - p(t + \tau)) \approx \sigma_p^2 e^{-\Lambda \tau}, \quad (20)$$

with the single measurement error of a two state system $\sigma_p^2(t) = \Phi(t)(1 - \Phi(t))$. We consider the N -dependent stochastic variable $\Phi_N(t)$, which fluctuates around the mean average value $\Phi(t)$ with a variance $\sigma_{\Phi_N}^2 = \sum_{n=0}^{N-1} \sigma_p^2(\partial \Phi_N / \partial p_n)^2 + \sum_{n \neq m}^{N-1} (\partial \Phi_N / \partial p_n)(\partial \Phi_N / \partial p_m) \text{cov}(p_n, p_m)$, where we omitted the t -dependence. Together with eq. (20) we get approximately a geometric sum, which yields

$$\sigma_{\Phi_N}^2(t) \simeq \frac{\sigma_p^2(t)}{N(1 - e^{-\Lambda T})}. \quad (21)$$

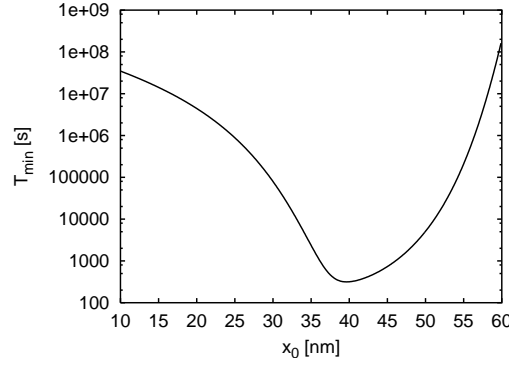


Fig. 5 – The minimum total measuring time T_{\min} over the offset extension x_0 for the optimized frequency $\omega^* \simeq \Lambda$ and the same parameters as in figs. 3 and 4.

The square root of the variance of this stochastic variable can now be identified with the measuring error, which should be smaller than the response amplitude Π_0 . Thus, we get the condition $\sqrt{\sigma_{\Phi_N}^2} \lesssim |\Pi_0|$. The total measuring time must therefore obey

$$T_{\text{tot}} \gtrsim \frac{2\pi\Phi_0(1-\Phi_0)(\omega^2 + \Lambda^2)}{\alpha^2\omega(1 - e^{-\Lambda 2\pi/\omega})}, \quad (22)$$

where we have replaced the phase sensitive $\Phi(t)(1-\Phi(t))$ by $\Phi_0(1-\Phi_0)$, which is permissible in zeroth order of $O(x_a/x_0)$. The right hand side of equation (22) has a minimum at $\omega^* \simeq \Lambda$, if the exponential correction in the denominator is neglected. The total measuring time for the optimized frequency ω^* is $T_{\min} \equiv T_{\text{tot}}(\omega^*) = 4\pi\kappa_{12}\kappa_{21}/(\Lambda\alpha^2)$, which strongly depends on the offset extension x_0 , see fig. 5. The minimum around x_0^* is a signature of a stochastic resonance-like phenomenon. Usually in stochastic resonance, the potential and a periodic driving input is given while the noise strength is controlled. One then finds an optimal noise strength for the best amplification of the input signal. In our case the strength of the noise is fixed by the temperature, whereas the driving force can easily be manipulated by adjusting offset x_0 , amplitude x_a and frequency ω . Figs. 3 and 5 show that where the response amplitude Π_0 as a function of offset x_0 is the largest the total measuring time T_{\min} is the smallest. Hence, the signal to noise ratio is the best for this choice of offset and optimized frequency. Moreover, it should be noted that typically in stochastic resonance the bistable potential is symmetric whereas in our case it is intrinsically asymmetric due to the molecular interactions.

In summary, we have shown that the two spontaneous transition rates k_{12}^0 , k_{21}^0 and the two most relevant length scales x_u and x_f of a biomolecular conformational transition can be extracted from data obtained by periodically driving these transitions. Our analysis focuses on the measured occupation probability of one state, which exhibits a stochastic resonance-like phenomenon as a function of the control parameters of a typical AFM experiment. Compared to more sophisticated approaches based on Jarzynski's relation, our scheme does not require to resolve thermal fluctuation within each minimum. Moreover, the linker molecule is explicitly included. Future applications of our scheme will include more complex transitions and linear sequences of similar transitions in the un- and refolding of multi-domain biopolymers. Most important, however, are experimental tests of periodic loading which we hope to encourage with this study.

* * *

We thank A. Hanke for fruitful interactions.

REFERENCES

- [1] MEHTA A. D., RIEF M., SPUDICH J. A., SMITH D.A. and SIMMONS R. M., *Science*, **283** (1999) 1689.
- [2] CLAUSEN-SCHAUMANN H., SEITZ M., KRAUTBAUER R. and GAUB H. E., *Curr. Opin. Chem. Biol.*, **4** (2000) 524.
- [3] EVANS E., *Ann. Rev. Biophys. Biomol. Struct.*, **30** (2001) 105.
- [4] MERKEL R., *Phys. Rep.*, **346** (2001) 344.
- [5] RIEF M. and GRUBMÜLLER H., *ChemPhysChem*, **3** (2002) 255.
- [6] STRICK T. R., DESSINGES M. N., CHARVIN G., DEKKER N. H., ALLEMAND J. F., BENSIMON D. and CROQUETTE V., *Rep. Prog. Phys.*, **66** (2003) 1.
- [7] BUSTAMANTE C., BRYANT Z. and SMITH S. B., *Nature*, **421** (2003) 423.
- [8] EVANS E. and RITCHIE K., *Biophys. J.*, **72** (1997) 1541.
- [9] JARZYNSKI C., *Phys. Rev. Lett.*, **78** (1997) 2690.
- [10] HUMMER G. and SZABO A., *Proc. Natl. Acad. Sci.*, **98** (2001) 3658.
- [11] LIPHARDT J., DUMONT S., SMITH S. B., TINOCO JR. I. and BUSTAMANTE C., *Science*, **296** (2002) 1832.
- [12] BRAUN O., HANKE A. and SEIFERT U., *cond-mat/0402496*, (2004) .
- [13] GAMMAITONI L., HÄNGGI P., JUNG P. and MARCHESONI F., *Rev. Mod. Phys.*, **70** (1998) 223.
- [14] LIPHARDT J., ONOA B., SMITH S. B., TINOCO JR. I. and BUSTAMANTE C., *Science*, **292** (2001) 733.
- [15] BELL G. I., *Science*, **200** (1978) 618.
- [16] RIEF M., FERNANDEZ J. M. and GAUB H. E., *Phys. Rev. Lett.*, **81** (1998) 4764.
- [17] MARKO J. F. and SIGGIA E. D., *Macromolecules*, **28** (1995) 8759.
- [18] OBERHAUSER A. F., MARSZALEK P. E., ERICKSON H. P. and FERNANDEZ J. M., *Nature*, **393** (1998) 181.